

## ABSTRACT

5        This invention comprises analyzing a speaker's speech in an audio visual  
recording to convert it into triphones and/or phonemes and then using a time coded  
phoneme stream to identify corresponding visual facial motions, to create single frame  
snapshots or multi-frame clips of facial motion corresponding to speech phoneme  
utterance states and transformations, which are stored in a database, and which are  
10 subsequently used to animate the original speaker's face, synchronized to a new voice  
track that has been converted into a time-coded, image frame-indexed phoneme stream.